

# Finding the Dark Hole with the Lights On: A New Approach to Focal Plane Wavefront Sensing

Richard A. Frazin\*

\*Dept. of Climate and Space Sciences, University of Michigan, Ann Arbor, MI 48109

## ABSTRACT

In direct imaging of exoplanets from space, achieving the required dynamic range (i.e., planet-to-star contrast in brightness) currently relies on coronagraphic technology combined with active control of one or more deformable mirrors (DMs) to create a dark region in the image plane, sometimes called a “dark hole.” While many algorithms have been proposed for this purpose, all of them employ focal plane wavefront sensing (FPWS) in order to calculate the optimal DM configuration to create the desired dark hole. All current algorithms are limited by their own success in that, as the dark hole is achieved, the FPWS procedure becomes shot-noise limited due to the low intensity in the dark hole. This article proposes a FPWS procedure that allows determination of the optimal DM configuration without relying on information obtained when the DM is near the optimal configuration. This article gives regression procedures for FPWS that do not assume the DM step size is small, which should allow two important improvements to the control loop: 1) performing informative FPWS observations with DM configurations that are sufficiently distant from the optimal dark hole configuration to mitigate shot-noise limitations, and 2) more accurately predicting the DM configuration that will achieve the desired objective in the dark hole control loop. In order to treat this more challenging FPWS problem, two different representations are presented. The first of these, is called the empirical Green’s function (EGF), is easy to implement, and has a block-diagonal matrix structure that is well-suited to parallel processing. The other representation, based on an explicit aberration expansion (EAE) requires the regression to estimate a smaller number of parameters than the EGF, but leads to a dense matrix structure. The EGF and EAE methods both simultaneously estimate the planetary image.

## 1. INTRODUCTION

Direct imaging of exoplanetary systems is difficult due to the high contrast in brightness between the planet and the star, which results in the planetary light being buried beneath details of telescope’s point spread function (PSF) that can change on a large range of time-scales, ranging from minutes to days. At this time, the scientific community is planning to perform direct imaging of exoplanets with telescopes in space, with the first such effort being NASA’s WFIRST mission, which will use coronagraphic optics to suppress the starlight [[http://wfirst.gsfc.nasa.gov/science/presentations/vugraphs/SDT\\_Jul\\_2012.pdf](http://wfirst.gsfc.nasa.gov/science/presentations/vugraphs/SDT_Jul_2012.pdf)]. Even with a coronagraph, time-variable aberrations in the optical system (mostly due to thermal stresses) are substantial enough to necessitate the use of active optical elements, namely deformable mirrors (DMs), to create a region in the image plane that is dark enough to meet the mission requirements. Many methods for using the DM to create the dark region, sometimes called a “dark hole,” have been proposed in the years following the initial idea, which is attributed to Malbet et al. in 1995.<sup>1</sup> In order to calculate the required DM deformation, all of these methods (reviewed in [2]) employ a focal plane wavefront sensing (FPWS) technique in which the DM implements a series “probe shapes” in order to determine the electric value in image plane. One fundamental limitation in this approach is that as the DM approaches the required to make a dark hole, the hole does indeed become quite dark, so much so that the measurements needed to make the hole still darker are shot-noise limited due to the small intensity. One way to overcome this limitation is to determine the electric field without relying on measurements in which the DM is in a dark hole configuration. This is not possible with current FPWS methods because they assume that they do not model the optical system sufficiently accurately to allow large DM steps and assume that the product of the phase induced by the DM step and the phase caused by unknown aberration is small.<sup>2</sup> The regression

---

Email: rfrazin at umich.edu

framework presented in this article models the unknown aberration in the optical system in such a way that this limitation on the step-size is removed, at least in theory (the remaining step-size limitation is DM calibration and reproducibility).

In the initial paper on the subject in 1995, Malbet et al. assumed that the unknown aberration in the optical system can be represented as equivalent aberration in a pupil plane upstream of the DM.<sup>1</sup> Unfortunately, is an oversimplification for optical systems that have aberration downstream of the DM. One optical system that is very likely to have aberrations downstream of the DM is the coronagraph for NASA’s WFIRST mission, the beam interacts with well over a dozen optical surfaces after bouncing off of “DM1,” so there is a high probability that the equivalent aberration will vary as a consequence of a probe command sent to DM1.<sup>3</sup>

This paper uses rigorous physical optics arguments to find the term required to represent the aberrations downstream of the DM and show that this term is of the same order as Malbet et al.’s aberration. After demonstrating the need for this additional term, regression equations are derived for two different representations of the post-DM aberrations, and their relative merits are presented. Finally, this article compares this FPWS approach to those that have been published and implemented previously.

## 2. PROPAGATION EQUATIONS

First, this section introduces some notation and formalism. The equations in this section will be within paradigm of scalar fields. Vector generalizations to handle polarizing optical systems of some equations are given by [4], but this important issue will be deferred to a sequel to this article.

Let  $\lambda$  be the central wavelength of the quasi-monochromatic light, and  $u(\mathbf{r})$  be the analytical signal representing the electric field [4–6] in some plane with two-dimensional (2D) coordinate vector  $\mathbf{r}$ . Note that, consistently with the formalism established in [4], the high-frequency factor  $\exp(-j2\pi\nu t)$  (where the speed of light  $c = \lambda\nu$ ) has been suppressed. The time-dependence due to statistical coherence properties of the field and will not be carried in the notation, either.

Consider an optical system on a space-based platform, such as in Fig. 1, in which the entrance pupil is designated as plane 0. Assuming the angular size of the star is small enough to allow treatment as a point source, the field arising from the star hosting the putative planetary impinging on the entrance pupil is [4]:

$$u_{\star 0}(\mathbf{r}_0) = \sqrt{I_\star} \exp(jk\boldsymbol{\alpha}_\star \cdot \mathbf{r}) \quad (1)$$

where  $k = 2\pi/\lambda$  and  $\boldsymbol{\alpha}_\star$  is the 2D sky angle, in units of radians, of the (presumably small) pointing error, and  $I_\star$  is the star’s irradiance. On the other hand, the entrance pupil is presumably large enough to at least partially resolve the planetary system, so it is therefore not possible to write a general expression for the planetary field impinging on the telescope entrance pupil (unless it is assumed to be composed of unresolved point sources). Instead, all that is available is the mutual coherence function, as provided by the well-known Van Cittert-Zernike theorem [4, 6]:

$$\gamma_{p0}(\mathbf{r}_0, \mathbf{r}'_0) = \frac{1}{\pi} \int_{\mathbf{p}} d\boldsymbol{\alpha} S(\boldsymbol{\alpha}) \exp[jk(\mathbf{r}_0 - \mathbf{r}'_0) \cdot \boldsymbol{\alpha}] , \quad (2)$$

where  $\boldsymbol{\alpha}$  is the 2D sky angle (radian units),  $S(\boldsymbol{\alpha})$  is the radiance of planetary system, and the integration is carried out over the angular extent of the planetary system, but not including the star itself (practically, this will be enforced by not trying to determine the image within some radius of the pointing center). Eq. (2) states that the mutual coherence of the light arriving at the Earth is proportional to the Fourier transform of the radiance emerging from the planetary system,  $S(\boldsymbol{\alpha})$ . The objective of the science observations is to estimate  $S(\boldsymbol{\alpha})$ , which is difficult because the stellar irradiance is much brighter than the planetary irradiance, i.e.,  $I_\star \gg \int_{\mathbf{p}} d\boldsymbol{\alpha} S(\boldsymbol{\alpha})$ .

The optical system shown schematically in Fig. 1 is intended to represent a space-based telescope as it does not include atmospheric modulation of light. It is assumed to contain only a single DM, as generalization to the multiple DM case is straightforward, at least at the theoretical level presented in this article. The light from the entrance pupil (whose plane is assigned the index number 0) passes through the part of the optical bench called the “pre-DM optics,” reflects off the DM, denoted with index  $d$ , then passes through “post-DM optics” (presumably including a coronagraph), finally forming an image on the science camera (SC), whose detector

surface will be given assigned index  $c$ . Let the operator  $\Upsilon_{d,0}(\mathbf{r}_d, \mathbf{r}_0)$  (where  $\mathbf{r}_d$  and  $\mathbf{r}_0$  are the 2D coordinates in the  $d$  and 0 planes, respectively) propagate the field from the telescope entrance pupil to just before the DM, thus including all effects of the pre-DM optics. The stellar field impinging on the DM is given by

$$u_d(\mathbf{r}_d) = \Upsilon_{d,0}(\mathbf{r}_d, \mathbf{r}_0) u_{\star 0}(\mathbf{r}_0), \quad (3)$$

in which integration over the  $\mathbf{r}_0$  coordinate is implied by the notation. Now, the DM imposes a phase shift of  $\zeta(\mathbf{r}_d, \mathbf{m})$  on the beam, where the second argument  $\mathbf{m}$  indicates that the DM shape ( $\zeta$ ) is function of the DM command vector  $\mathbf{m}$ . Here it will be assumed that  $\mathbf{m}$  changes on a time-scale that is much shorter than  $\tau_d$ , the dynamical time-scale on which the spacecraft structure undergoes thermal relaxations and so on. In this treatment, changes on the time-scale  $\tau_d$  are not included and quantities that vary on that time-scale are formally considered to be constant. As a practical matter, in order to deal with changes on the time-scale  $\tau_d$ , the regressions presented below will need to be updated regularly, most likely in a Kalman-filtering framework.<sup>7</sup>

Here, we will assume that  $\zeta(\mathbf{r}_d, \mathbf{m})$  is a purely real-valued, which does not allow the DM to have amplitude effects (lifting this restriction is relatively straightforward). The field just after reflecting off the DM is then given by

$$u_d^+(\mathbf{r}_d, \mathbf{m}) = u_d(\mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] = \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}(\mathbf{r}_d, \mathbf{r}_0) u_{\star 0}(\mathbf{r}_0), \quad (4)$$

where the  $^+$  superscript emphasizes that the light has just reflected off of DM. Similarly,  $\Upsilon_{c,d}(\mathbf{r}_c, \mathbf{r}_d)$ , where  $\mathbf{r}_c$  is the 2D coordinate in the SC plane, propagates  $u_d^+$  from the DM to the SC. The propagator  $\Upsilon_{c,d}$  includes the post-DM optical train. Note, in a break with the conventions established in [4],  $\Upsilon_{c,d}$  does not include interaction with the surface specified by index  $d$ , which, in this case, is taken into account by the factor  $\exp[j\zeta(\mathbf{r}_d, \mathbf{m})]$ . The field in the SC arising from the starlight is given by

$$\begin{aligned} u_c(\mathbf{r}_c, \mathbf{m}) &= \Upsilon_{c,d}(\mathbf{r}_c, \mathbf{r}_d) u_d^+(\mathbf{r}_d) \\ &= \Upsilon_{c,d}(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}(\mathbf{r}_d, \mathbf{r}_0) u_{\star 0}(\mathbf{r}_0). \end{aligned} \quad (5)$$

Under this formalism, Eq. (5) implies integration not only over the pupil plane coordinate  $\mathbf{r}_0$ , but also over the intermediate coordinate in the DM plane,  $\mathbf{r}_d$ .

Due to unknown and time-variable (on the  $\tau_d$  time-scale) aberrations throughout the entire optical system, the propagators  $\Upsilon_{c,d}$  and  $\Upsilon_{d,0}$  are only partially known. It is helpful to decompose  $\Upsilon_{c,d}$  and  $\Upsilon_{d,0}$  into known (denoted with superscript  $k$ ) and unknown (superscript  $u$ ) as follows:

$$\Upsilon_{c,d}(\mathbf{r}_c, \mathbf{r}_d) = \Upsilon_{c,d}^k(\mathbf{r}_c, \mathbf{r}_d) + \Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \quad (6)$$

$$\Upsilon_{d,0}(\mathbf{r}_d, \mathbf{r}_0) = \Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0) + \Upsilon_{d,0}^u(\mathbf{r}_d, \mathbf{r}_0), \quad (7)$$

in which the known propagators  $\Upsilon_{d,0}^k$  and  $\Upsilon_{c,d}^k$  must be implemented with numerical integrations. The unknown propagators  $\Upsilon_{c,d}^u$  and  $\Upsilon_{d,0}^u$  account for the unknown aberrations in the post-DM and pre-DM optical trains, respectively. Using Eqs. (6) and (7) in Eq. (5) results in an expression for the field consisting of known and unknown parts (using the same superscripting convention), i.e.,

$$u_c(\mathbf{r}_c, \mathbf{m}) = u_c^k(\mathbf{r}_c, \mathbf{m}) + u_c^u(\mathbf{r}_c, \mathbf{m}), \quad (8)$$

where

$$u_c^k(\mathbf{r}_c, \mathbf{m}) = \Upsilon_{c,d}^k(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0) u_{\star 0}(\mathbf{r}_0), \quad (9)$$

and

$$\begin{aligned} u_c^u(\mathbf{r}_c, \mathbf{m}) &= \{ \Upsilon_{c,d}^k(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^u(\mathbf{r}_d, \mathbf{r}_0) + \\ &\quad \Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0) + \Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^u(\mathbf{r}_d, \mathbf{r}_0) \} u_{\star 0}(\mathbf{r}_0) \end{aligned} \quad (10)$$

It is helpful to define the unknown quantity  $g^u(\mathbf{r}_d)$ , called the *effective pre-DM aberration*:

$$g^u(\mathbf{r}_d) \equiv \Upsilon_{d,0}^u(\mathbf{r}_d, \mathbf{r}_0) u_{\star 0}(\mathbf{r}_0). \quad (11)$$

Eq. (11) states that all of the aberrations in the pre-DM optics result in an unknown function of the coordinate in the DM plane,  $g^u(\mathbf{r}_d)$ , and is essentially the same as the aberration originally assumed by Malbet et al.,<sup>1</sup> as will be explained below in Sec. 5. Note that  $g^u(\mathbf{r}_d)$  also includes any unknown aberration in the DM itself that varies on the time-scale  $\tau_d$ , which, unfortunately excludes uncalibrated deformations of the DM shape that vary as its configuration changes. Using Eq. (11) to simplify Eq. (10), one obtains:

$$u_c^u(\mathbf{r}_c, \mathbf{m}) = \Upsilon_{c,d}^k(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})]g^u(\mathbf{r}_d) + \Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})]\Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0)u_{*0}(\mathbf{r}_0) + \Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})]g^u(\mathbf{r}_d). \quad (12)$$

One can see that  $g^u$  and  $\Upsilon_{c,d}^u$  should be of roughly equal importance, as the aberrated field  $u_c$  in Eq. (12) has terms that linear in each one, as well as a final term that bi-linear in the two quantities. Below, in Sec. 4.2, it will be seen that the importance of the  $\Upsilon_{c,d}^u$  term should be proportional to the number optical surfaces in the post-DM optical train (or at least the number of fully illuminated surfaces that are upstream of the coronagraph).

### 3. SCIENCE CAMERA INTENSITY

The intensity of the light impinging on the SC is given by the sum of the planetary intensity and the stellar intensity:

$$I_c(\boldsymbol{\rho}, \mathbf{m}) = I_{pc}(\boldsymbol{\rho}, \mathbf{m}) + I_c(\boldsymbol{\rho}, \mathbf{m}) \quad (13)$$

where  $\boldsymbol{\rho} \equiv \mathbf{r}_c$  is the coordinate in the SC plane. Now, the planetary intensity  $I_{pc}$  is vastly fainter than the stellar intensity and a relatively simple model of the optical system, with no unknown aberration, should be perfectly adequate for calculating the planetary contribution. Let the needed simplified propagation operator be denoted by  $\Upsilon_p(\boldsymbol{\rho}, \mathbf{r}_0, \mathbf{m})$ . Choosing  $\Upsilon_p(\boldsymbol{\rho}, \mathbf{r}_0, \mathbf{m}) = \Upsilon_{c,d}^k(\boldsymbol{\rho}, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})]\Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0)$  should certainly be a valid option, but  $\Upsilon_{c,d}^k$  and  $\Upsilon_{d,0}^k$  operators may be unnecessarily computationally expensive, so that cheaper approximations may serve for calculating the planetary contribution. Any approximation should include the effect of the DM ( $\exp[j\zeta(\mathbf{r}_d, \mathbf{m})]$ ), as it will modulate the planetary light. In any event, the planetary intensity on the SC is given by integrating over the mutual coherence function in Eq. (2) [4, 6]:

$$\begin{aligned} I_{pc}(\boldsymbol{\rho}, \mathbf{m}) &= \Upsilon_p(\boldsymbol{\rho}, \mathbf{r}_0, \mathbf{m})\Upsilon_p^*(\boldsymbol{\rho}, \mathbf{r}'_0, \mathbf{m})\gamma_{p0}(\mathbf{r}_0, \mathbf{r}'_0) \\ &= \Upsilon_p(\boldsymbol{\rho}, \mathbf{r}_0, \mathbf{m})\Upsilon_p^*(\boldsymbol{\rho}, \mathbf{r}'_0, \mathbf{m})\frac{1}{\pi}\int_p d\boldsymbol{\alpha} S(\boldsymbol{\alpha}) \exp[jk(\mathbf{r}_0 - \mathbf{r}'_0) \cdot \boldsymbol{\alpha}] \\ &= \frac{1}{\pi}\int_p d\boldsymbol{\alpha} S(\boldsymbol{\alpha}) \left\{ \Upsilon_p(\boldsymbol{\rho}, \mathbf{r}_0, \mathbf{m})\Upsilon_p^*(\boldsymbol{\rho}, \mathbf{r}'_0, \mathbf{m}) \exp[jk(\mathbf{r}_0 - \mathbf{r}'_0) \cdot \boldsymbol{\alpha}] \right\}, \end{aligned} \quad (14)$$

where the reader is reminded that that Eq. (14) is, in fact, a triple integral as the propagation operators imply integration over  $\mathbf{r}_0$  and  $\mathbf{r}'_0$ . The quantity inside the braces is called the *planetary intensity kernel* [4, 8]. In the absence of the confounding starlight, estimating  $S(\boldsymbol{\alpha})$  amounts to solving a multi-frame deconvolution problem. In this deconvolution problem, the coronagraphic optics make the point-spread function (PSF) spatially variant, at least close to the inner working angle, and the PSF depends on the DM command vector  $\mathbf{m}$ , as well.

The stellar contribution to the intensity is given by  $I_c(\boldsymbol{\rho}, \mathbf{m}) = u_c(\boldsymbol{\rho}, \mathbf{m})u_c^*(\boldsymbol{\rho}, \mathbf{m})$ , using Eq. (8) it can be written as:

$$I_c(\boldsymbol{\rho}, \mathbf{m}) - I_0(\boldsymbol{\rho}, \mathbf{m}) = u_c^k(\boldsymbol{\rho}, \mathbf{m})u_c^{u*}(\boldsymbol{\rho}, \mathbf{m}) + u_c^u(\boldsymbol{\rho}, \mathbf{m})u_c^{k*}(\boldsymbol{\rho}, \mathbf{m}) + u_c^u(\boldsymbol{\rho}, \mathbf{m})u_c^{u*}(\boldsymbol{\rho}, \mathbf{m}), \quad (15)$$

where the known portion of the intensity,  $I_0(\boldsymbol{\rho}, \mathbf{m}) \equiv u_c^k(\boldsymbol{\rho}, \mathbf{m})u_c^{k*}(\boldsymbol{\rho}, \mathbf{m})$ , has been moved to left-hand-side of the equation. The small, unknown aberrations in the optical system are taken into account by the unknown function  $g^u$  and the unknown operator  $\Upsilon_{c,d}^u$ . Without defining precise metrics to be more quantitative,  $g^u$ , defined in Eq. (11), is presumably smaller than  $\Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0)u_{*0}(\mathbf{r}_0)$  and  $\Upsilon_{c,d}^u$  should be smaller than  $\Upsilon_{c,d}^k$ , so that  $u_c^u$  in Eq. (12) contains terms that are of 1st and 2nd order in these unknown quantities. It then follows that the right-hand-side of Eq. (15) contains terms that of 1st, 2nd, 3rd and 4th order in these small, unknown quantities.

Below, only the 1<sup>st</sup> order terms are given, as deriving the others is straightforward, and they can be treated by standard procedures involving linearization and iteration, if necessary. Then,

$$I_c(\boldsymbol{\rho}, \mathbf{m}) - I_0(\boldsymbol{\rho}, \mathbf{m}) \approx u_c^k(\boldsymbol{\rho}, \mathbf{m}) \{ \Upsilon_{c,d}^{k*}(\mathbf{r}_c, \mathbf{r}_d) \exp[-j\zeta(\mathbf{r}_d, \mathbf{m})] g^{u*}(\mathbf{r}_d) + \Upsilon_{c,d}^{u*}(\mathbf{r}_c, \mathbf{r}_d) \exp[-j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^{k*}(\mathbf{r}_d, \mathbf{r}_0) u_{*0}^*(\mathbf{r}_0) \} + u_c^{k*}(\boldsymbol{\rho}, \mathbf{m}) \{ \Upsilon_{c,d}^k(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] g^u(\mathbf{r}_d) + \Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0) u_{*0}(\mathbf{r}_0) \}, \quad (16)$$

where the reader will recall that the phase imparted by the DM,  $\zeta(\mathbf{r}_d, \mathbf{m})$ , was taken to be real.

The science camera measurement is subject to the effects of noise, both due to photon counting statistics [6] (shot noise) and readout noise, and perhaps other effects. Fortunately, there is a new generation of ultra-low noise IR cameras capable of kHz readouts, such as the SWIR single photon detector, SAPHIRA eAPD and the MKIDS,<sup>9–12</sup> making the issue of readout noise far less critical. Taking the noise into account, the measured value of the science camera intensity is obtained by summing the stellar and planetary contributions:

$$I_m(\boldsymbol{\rho}_l, \mathbf{m}_i) = (\Delta t_i) I_c(\boldsymbol{\rho}_l, \mathbf{m}_i) + (\Delta t_i) I_{pc}(\boldsymbol{\rho}_l, \mathbf{m}_i) + \nu(\boldsymbol{\rho}_l, t_i), \quad (17)$$

where  $\boldsymbol{\rho}_l$  is the position of the  $l$ <sup>th</sup> SC detector pixel,  $\mathbf{m}_i$  is the  $i$ <sup>th</sup> DM command vector,  $t_i$  is the time-stamp associated with  $\mathbf{m}_i$ ,  $\Delta t_i$  is the corresponding exposure time, and  $\nu(\boldsymbol{\rho}_l, t_i)$  is a sample of the random process that describes the detector noise. Eq. (17) makes several assumptions:

- The detector pixels are small enough so that integration of  $I_c(\boldsymbol{\rho}, \mathbf{m})$  and  $I_{pc}(\boldsymbol{\rho}, \mathbf{m})$  over the pixel area is not necessary. Implementing such an integration is straightforward.
- The noise,  $\nu(\boldsymbol{\rho}_l, t_i)$  has no important dependence on the unknown quantities  $S(\boldsymbol{\alpha})$ ,  $g^u(\mathbf{r}_d)$  and the unknown operator  $\Upsilon_{c,d}^u(\boldsymbol{\rho}_l, \mathbf{r}_d)$ . This cannot be strictly true as the shot noise will depend on these values, especially as the success is achieved in creating a dark hole. In an iterative procedure, the statistics of  $\nu$  can be calculated with current estimates of unknowns, if necessary. However, the entire point of the FPWS strategy presented here is to gain the required information from DM configurations in which shot noise is not a big problem in the first place.

#### 4. REGRESSION STRATEGIES

The regression procedure estimates the effective pre-DM aberration  $g^u$ , the post-DM aberrations  $\Upsilon_{c,d}^u$  and the planetary image  $S$ . Most of the information pertaining to  $g^u$  and  $\Upsilon_{c,d}^u$  comes from instances when the DM is not making a dark hole, while the estimate of  $S$  depends most heavily on the instances when the hole is darkest. The objective of this section is to give the regression formulae for estimating these quantities.

Since  $S(\boldsymbol{\alpha})$  and  $g^u(\mathbf{r}_d)$  are continuous functions on some portion of  $\mathbb{R}^2$  and  $\mathbb{C}^2$ , respectively, they must be parameterized somehow for computation. The most natural way to achieve this is via series expansions, which reduces the unknown continuous quantities to a finite set of expansion coefficients. The image processing community has a vast literature on how best to do this, including multi-resolution representations (e.g., wavelets), but these issues will not be treated here, and instead, only generic forms will be given. Estimating the kernel of the propagation operator  $\Upsilon_{c,d}^u(\boldsymbol{\rho}_l, \mathbf{r}_d)$  is more challenging because it defines a continuous mapping  $\Upsilon_{c,d}^u : \mathbb{C}^2 \rightarrow \mathbb{R}^2$ . Below, two strategies for estimating  $\Upsilon_{c,d}^u$  will be presented. The first approach is called the *Empirical Green's Function*, and despite its high dimensionality, it is (relatively) straightforward to implement and very well-suited to parallel computation. The EGF was first introduced by the author in [13] in the context of exoplanet imaging from ground-based observatories. The second approach, called *Explicit Aberration Expansion*, requires challenging and expensive Fresnel propagation computations and is much less easy to parallelize. However, in most situations, it will likely require the estimation of many fewer unknown quantities than the EGF and therefore should make more efficient use of observational resources, at least in theory.

The regression can be expressed in linear algebra terms via the canonical equation  $\mathbf{y} = \mathbf{H}\mathbf{x}$ , where  $\mathbf{y}$  is derived from the observations (SC pixel values),  $\mathbf{x}$  is a vector of unknown regression coefficients to be estimated, and  $\mathbf{H}$

is the model-based matrix that relates the two. Generalization to a Kalman filtering framework, as in [7], is left to future work. The vector  $\mathbf{x}$  is composed of three sub-vectors as follows:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_\Upsilon \\ \mathbf{x}_g \\ \mathbf{x}_p \end{bmatrix}, \quad (18)$$

where  $\mathbf{x}_\Upsilon$  contains the regression coefficients associated with estimating  $\Upsilon_{c,d}^u$ , and  $\mathbf{x}_g$  and  $\mathbf{x}_p$  contain the coefficients specifying the effective pre-DM aberration  $g^u$  and the planetary image  $S$ , respectively. Correspondingly, the matrix  $\mathbf{H}$  is partitioned into three submatrices as:

$$\mathbf{H} = [\mathbf{H}_\Upsilon \quad \mathbf{H}_g \quad \mathbf{H}_p], \quad (19)$$

where the subscripts have meanings that correspond those in Eq. (19).

The vector  $\mathbf{y}$  is modeled by Eq. (17), and is arranged so that each subvector  $\mathbf{y}_l$  corresponds to the time-series associated with the  $l$ th SC pixel. Recalling that the time-stamp  $t_i$  corresponds to the DM command  $\mathbf{m}_i$  by given by:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \vdots \\ \mathbf{y}_l \\ \vdots \\ \mathbf{y}_{N-1} \end{bmatrix}, \quad (20)$$

where  $N$  is the number of SC pixels and

$$\mathbf{y}_l = \begin{bmatrix} I_m(\boldsymbol{\rho}_l, \mathbf{m}_0) - I_0(\boldsymbol{\rho}_l, \mathbf{m}_0) - \nu(\boldsymbol{\rho}_l, t_0) \\ \vdots \\ I_m(\boldsymbol{\rho}_l, \mathbf{m}_i) - I_0(\boldsymbol{\rho}_l, \mathbf{m}_i) - \nu(\boldsymbol{\rho}_l, t_i) \\ \vdots \\ I_m(\boldsymbol{\rho}_l, \mathbf{m}_{T-1}) - I_0(\boldsymbol{\rho}_l, \mathbf{m}_{T-1}) - \nu(\boldsymbol{\rho}_l, t_{T-1}) \end{bmatrix}, \quad (21)$$

in which  $T$  is the number of DM steps used in the estimation. Thus,  $\mathbf{y}_l$  has  $T$  elements and  $\mathbf{y}$  has  $NT$  elements. The unknown effective pre-DM aberration  $g^u(\mathbf{r})$  can be approximated with the following expansion:

$$g^u(\mathbf{r}) \approx \sum_{l=0}^{N_g-1} (a_l + jb_l) \psi_l^g(\mathbf{r}), \quad (22)$$

where  $N_g$  is the number of terms, and the  $\{\psi_l^g(\mathbf{r})\}$  are real-valued expansion functions (e.g., annular Zernike polynomials). Since  $g^u(\mathbf{r})$  is complex-valued, the expansion has a real part given by the expansion coefficients  $\{a_l\}$ , and the imaginary part is given by the expansion coefficients  $\{b_l\}$ . The vector  $\mathbf{x}_g$  is given by:

$$\mathbf{x}_g^T = [a_0, \dots, a_l, \dots, a_{N_g-1}, b_0, \dots, b_l, \dots, b_{N_g-1}]^T, \quad (23)$$

where the superscript  $T$  indicates tranposition (in this case, to make it into a column vector). The corresponding elements of the  $NT \times 2N_g$  matrix  $\mathbf{H}_g$  are calculated by inserting Eq. (22) into Eq. (16) and performing the requisite integrations for each DM position  $\mathbf{m}_i$  and SC pixel position  $\boldsymbol{\rho}_l$ . Similarly, the planetary image  $S(\boldsymbol{\alpha})$  can be approximated with a series expansion as follows:

$$S(\boldsymbol{\alpha}) \approx \sum_{l=0}^{N_p-1} p_l \psi_l^P(\boldsymbol{\alpha}), \quad (24)$$



where  $N_p$  is the number of terms included in the expansion, the  $\{p_l\}$  are the real-valued expansion coefficients, and the  $\{\psi_l^p(\boldsymbol{\alpha})\}$  are the real-valued expansion functions used for describing the planetary image. The vector  $\mathbf{x}_p$  in Eq. (19) holds the planetary coefficients and is given by:

$$\mathbf{x}_p^T = [p_0, \dots, p_l, \dots, p_{N_p-1}]^T. \quad (25)$$

The corresponding elements of the  $NT \times N_p$  matrix  $\mathbf{H}_p$  are calculated by inserting Eq. (24) into Eq. (14) and performing the needed integrations.

#### 4.1 Empirical Green's Function

The operator  $\Upsilon_{c,d}(\boldsymbol{\rho}, \mathbf{r})$  is Green's function that solves diffraction problem with the field propagating from sources in plane  $d$  to plane  $c$ .<sup>5</sup>  $\Upsilon_{c,d}^k(\boldsymbol{\rho}, \mathbf{r})$  is the known part of this operator, and the part that must be determined from the measurements,  $\Upsilon_{c,d}^u(\boldsymbol{\rho}, \mathbf{r})$ , will be called the Empirical Green's Function (EGF). When treating polarization effects, it has a straightforward generalization to vector fields, called the *Empirical Green's Tensor (EGT)*, which will be explained in the followup to this article [it was introduced in 13].

At any fixed position in the detector plane,  $\boldsymbol{\rho}_l$  the kernel of the operator  $\Upsilon_{c,d}^u(\boldsymbol{\rho}_l, \mathbf{r})$  is function on some region of  $\mathbb{R}^2$  (as is  $g^u(\mathbf{r})$ ). Then, at any given  $\boldsymbol{\rho}_l$ , it also admits a series expansion:

$$\Upsilon_{c,d}^u(\boldsymbol{\rho}_l, \mathbf{r}) \approx \sum_{k=0}^{M_l-1} (c_{l,k} + jd_{l,k}) \psi_k^l(\mathbf{r}), \quad (26)$$

where  $M_l$  is the number of terms in the series corresponding to the  $l$ th detector pixel, the  $\{\psi_k^l\}$  are the expansion functions corresponding to the  $l$ th detector pixel. Eq. (26) is called the EGF expansion. Since  $\Upsilon_{c,d}^u(\boldsymbol{\rho}_l, \mathbf{r})$  is complex-valued, the expansion has a real part given by the expansion coefficients  $\{c_{l,k}\}$  and the imaginary part is given by the expansion coefficients  $\{d_{l,k}\}$ . It is important to emphasize that, from this point of view,  $\Upsilon_{c,d}^u(\boldsymbol{\rho}_l, \mathbf{r})$  and  $\Upsilon_{c,d}^u(\boldsymbol{\rho}_m, \mathbf{r})$  are completely independent functions of  $\mathbf{r}$  unless  $l = m$ , so that the choices made for the  $\{\psi_k^l\}$  (as well as  $M_l$ ) can be made independently for each SC pixel position. It is, in fact, this independence that makes implementation of the EGF so easy to parallelize, as one obtains an independent regression for each pixel position  $l$ . This independence corresponds to a block-diagonal structure in linear algebra terms, as will be demonstrated directly. The vector  $\mathbf{x}_\Upsilon$  associated with the EGF is composed of  $N$  (one for each SC pixel) sub-vectors as follows:

$$\mathbf{x}_\Upsilon^T = [\mathbf{x}_0, \dots, \mathbf{x}_l, \dots, \mathbf{x}_{N-1}]^T, \quad (27)$$

in which

$$\mathbf{x}_l = [c_{l,0}, \dots, c_{l,k}, \dots, c_{l,M_l-1}, d_{l,0}, \dots, d_{l,k}, \dots, d_{l,M_l-1}]^T, \quad (28)$$

which as  $2M_l$  components. By creating an entirely separate expansion series for each SC pixel, Eq. (26) assures that the matrix  $\mathbf{H}_\Upsilon$  will have a block-diagonal representation. Let  $\mathbf{H}_{\Upsilon,l}$  be the  $T \times 2M_l$  block corresponding to the sub-vector  $\mathbf{x}_l$  in Eq. (28). Its elements are calculated by inserting Eq. (26) into Eq. (16). The resulting linear system has the form:

$$\begin{bmatrix} \mathbf{y}_0 \\ \vdots \\ \mathbf{y}_l \\ \vdots \\ \mathbf{y}_{N-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{\Upsilon,0} & & & \mathbf{H}_{g,0} & \mathbf{H}_{p,0} \\ & \ddots & & \vdots & \vdots \\ & & \mathbf{H}_{\Upsilon,l} & \mathbf{H}_{g,l} & \mathbf{H}_{p,l} \\ & & & \ddots & \vdots \\ & & & & \mathbf{H}_{\Upsilon,N-1} & \mathbf{H}_{g,N-1} & \mathbf{H}_{p,N-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_\Upsilon \\ \mathbf{x}_g \\ \mathbf{x}_p \end{bmatrix}, \quad (29)$$

in which  $\mathbf{H}_{g,l}$  is the portion of the  $\mathbf{H}_g$  matrix associated with the  $l$ th SC pixel, and similarly for  $\mathbf{H}_{p,l}$ .

The partial block-diagonal structure in Eq. (29) immediately suggests an iterative, alternating minimization procedure in which  $\mathbf{x}_g$  and  $\mathbf{x}_p$  are held constant while estimate of  $\mathbf{x}_\Upsilon$  is updated, and vice-versa (perhaps using algorithms that enjoy improved convergence rates by employing “costs-to-move”). Indeed, the planetary coefficients  $\mathbf{x}_p$  can probably be assumed to 0, except in pixels that happen to be inside a dark hole.

The price to be paid for the “parallelizability” and relative simplicity of the EGF approach is the fact specifying the EGF requires estimating  $2M_l$  free parameters for each SC pixel. It may well be practical to solve for a sufficient number of coefficients to allow a rather high-order expansion in Eq. (26), especially if high-cadence detectors are available. As an example, assume the expansion functions  $\{\psi_k^l(\mathbf{r})\}$  are given by the annular Zernike polynomials up to 20th order, which would correspond to a total of 230 polynomials, or  $M_l = 230$ . Since we must solve for the real and imaginary parts of the expansion coefficients, the vector  $\mathbf{x}_l$  would have 460 components that must be estimated from the regression. As a rather arbitrary example, assume that the cadence of the SC is 0.01 seconds, and that each exposure corresponds to a different DM command  $\mathbf{m}$ . After 100 seconds, each SC pixel would have  $10^4$  observations from which to estimate these 460 coefficients, overdetermining the problem by a factor of about 20. The number of coefficients that one may determine is likely to be limited by the inversion of a matrix of size  $2M_l \times 2M_l$  (computation time proportional to  $M_l^3$ ), or  $460 \times 460$  in this example. Note that the author’s desktop machine was able to invert a  $1000 \times 1000$  matrix of random numbers (which tends to result in a poorly conditioned system) in about 0.1 s.

## 4.2 Explicit Aberration Expansion

The explicit aberration expansion (EAE), in which the propagator  $\Upsilon_{c,d}^u(\boldsymbol{\rho}, \mathbf{r})$  is assumed to be the result of unknown aberration functions in  $P$  planes. The EAE was introduced in [4] in the context of ground-based imaging. The fact that the EAE should require estimating fewer parameters than the EGF can be seen as follows: In the EAE, each aberration function requires  $M_l = M$  expansion coefficients to be estimated, then there would be a total of  $PM$  in the  $\mathbf{x}_\Upsilon$  vector, whereas, in the EGF, the  $\mathbf{x}_\Upsilon$  would contain  $NM$  elements. One would expect  $PM \ll NM$  since the number of planes  $P$  is much smaller than the number of SC pixels  $N$ . Furthermore, in some optical systems the needed value of  $P$  can be reduced by treating aberrations in conjugate planes an equivalent aberration in one plane, as will be shown below. The disadvantages of the EAE relative to the EGF are:

- the first order approximation, which prohibits aberrated fields interacting with other aberrations (much like the Born approximation in scattering theory), whereas the EGF assumes only that the optical system is linear in the field, not the aberrations
- the difficult Fresnel integrations needed to propagate the field from one plane to the next
- the fact that the matrix  $\mathbf{H}_\Upsilon$  is dense, instead of block-diagonal, as it is in the EGF.

To derive the EAE, assume that the post-DM optics contain unknown aberrations in  $P$  planes, with the first plane in the post-DM optics containing aberration given the index 1 (the index 0 was already used to signify the telescope entrance pupil), and the final aberrated plane having the index  $P$ . Let the propagator between plane  $l$  and plane  $l+1$  be denoted by  $\Upsilon_{l+1,l}(\mathbf{r}_{l+1}, \mathbf{r}_l)$ , which includes interaction with the  $l$  surface but not the  $l+1$  surface. The only exception to this rule are propagators that start at the DM plane  $d$  (note that the is already included in  $g^u$ ), so  $\Upsilon_{1,d}(\mathbf{r}_1, \mathbf{r}_d) = \Upsilon_{1,d}^k(\mathbf{r}_1, \mathbf{r}_d)$  is a known operator.

Following the earlier procedure, the propagator  $\Upsilon_{l+1,l}$  can be decomposed into known and unknown parts:

$$\begin{aligned} \Upsilon_{l+1,l}(\mathbf{r}_{l+1}, \mathbf{r}_l) &= \Upsilon_{l+1,l}^k(\mathbf{r}_{l+1}, \mathbf{r}_l) + \Upsilon_{l+1,l}^u(\mathbf{r}_{l+1}, \mathbf{r}_l) \\ &= \Upsilon_{l+1,l}^k(\mathbf{r}_{l+1}, \mathbf{r}_l) [1 + A_l^u(\mathbf{r}_l)], \end{aligned} \quad (30)$$

which assumes that the unknown character of the propagator is confined to an unknown aberration function in plane  $l$ , given by the complex-valued function  $A_l^u(\mathbf{r}_l)$ . While it may be tempting to set  $A_l^u(\mathbf{r}_l) = \exp[j\phi_l(\mathbf{r}_l)]$  for some (possibly complex-valued) phase aberration function  $\phi_l(\mathbf{r}_l)$  as was done in [4], it is of little practical value as one must immediately perform the Taylor expansion of the exponential  $A_l^u(\mathbf{r}_l) \approx 1 + j\phi_l(\mathbf{r}_l) - \phi_l^2(\mathbf{r}_l)/2 + \dots$  and deal with consequences of the approximation. On the other hand, this problem is avoided by treating  $A_l^u$  as the sum of two functions with real and imaginary parts, i.e.,  $A_l^u(\mathbf{r}_l) = A_l^{ur}(\mathbf{r}_l) + jA_l^{ui}(\mathbf{r}_l)$ . The case of a small, real-valued phase perturbation  $\phi_l(\mathbf{r}_l)$  is recovered by fixing the value  $A_l^{ur}(\mathbf{r}_l) = 1$  and setting  $A_l^{ui}(\mathbf{r}_l) = \phi_l(\mathbf{r}_l)$ .

The known field leaving the DM, i.e., excluding the pre-DM aberrations is:

$$u_d^k(\mathbf{r}_d, \mathbf{m}) = \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0) u_{*0}(\mathbf{r}_0), \quad (31)$$



and using Eq. (30), the contribution to the aberrated field,  $u_C^u$ , arising from the aberration in plane  $l$  arriving at SC is:

$$\Upsilon_{c,l+1}^k(\mathbf{r}_c, \mathbf{r}_{l+1}) \Upsilon_{l+1,l}^k(\mathbf{r}_{l+1}, \mathbf{r}_l) A_l^u(\mathbf{r}_l) \Upsilon_{l,d}^k(\mathbf{r}_l, \mathbf{r}_d) u_d^k(\mathbf{r}_d, \mathbf{m}), \quad (32)$$

where  $\Upsilon_{l,d}^k$  is a known operator that propagates the field from the  $d$  plane to the  $l$  plane that is given by contraction:

$$\Upsilon_{l,d}^k(\mathbf{r}_l, \mathbf{r}_d) = \left[ \prod_{k=1}^{l-1} \Upsilon_{k+1,d}^k(\mathbf{r}_{k+1}, \mathbf{r}_l) \right] \Upsilon_{1,d}^k(\mathbf{r}_1, \mathbf{r}_d). \quad (33)$$

Similarly,  $\Upsilon_{c,l}^k(\mathbf{r}_c, \mathbf{r}_l) = \Upsilon_{c,l+1}^k(\mathbf{r}_c, \mathbf{r}_{l+1}) \Upsilon_{l+1,l}^k(\mathbf{r}_{l+1}, \mathbf{r}_l)$ , so that total contribution of the post-DM aberrations to the SC field is, to 1<sup>st</sup> order in the aberrations,

$$\sum_{l=1}^P \Upsilon_{c,l}^k(\mathbf{r}_c, \mathbf{r}_l) A_l^u(\mathbf{r}_l) \Upsilon_{l,d}^k(\mathbf{r}_l, \mathbf{r}_d) u_d^k(\mathbf{r}_d, \mathbf{m}). \quad (34)$$

Thus, the unknown propagator  $\Upsilon_{c,d}^u$  is given by:

$$\Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) \approx \sum_{l=1}^P \Upsilon_{c,l}^k(\mathbf{r}_c, \mathbf{r}_l) A_l^u(\mathbf{r}_l) \Upsilon_{l,d}^k(\mathbf{r}_l, \mathbf{r}_d), \quad (35)$$

where the strict equality is lost due to the fact that terms in which aberrated fields do not interact with downstream aberrations are excluded. This should be permissible if  $|A_l^u(\mathbf{r}_l)| \ll 1$ , as per Eq. (30). Eq. (35) is the EAE. One obvious consequence is that the value  $\Upsilon_{c,d}^u$  is roughly proportional to  $P$ , the number of planes containing aberration, which is likely every optical surface in the post-DM optical train, or at least the number of fully illuminated planes upstream of the coronagraph. Unlike the EGF, which has completely independent expansion for each SC position  $\mathbf{r}_c$ , in the EAE, each aberration  $A_l$  potentially influences the value of the field at each SC position. Note that implementation of the propagators  $\Upsilon_{c,l}^k(\mathbf{r}_c, \mathbf{r}_l)$  and  $\Upsilon_{l,d}^k(\mathbf{r}_l, \mathbf{r}_d)$  in Eq. (35) will likely require Fresnel integrations, which can be expensive and challenging.

The function  $A_l^u(\mathbf{r}_l)$  can be reduced to a finite parameter set by again employing a series expansion:

$$A_l^u(\mathbf{r}_l) \approx \sum_{k=0}^{M_l-1} (c_{l,k} + j d_{l,k}) \psi_k^l(\mathbf{r}_l), \quad (36)$$

where  $M_l$  is the number of terms in the sum, the coefficients  $\{c_{l,k}\}$  and  $\{d_{l,k}\}$  are real and the  $\{\psi_k^l(\mathbf{r}_l)\}$  are expansion functions. Eq. (36) looks rather similar to Eq. (26), but, in this case, the expansion approximates the aberration function in the  $l$ <sup>th</sup> plane. As mentioned above, if one is only concerned with small, real-valued phase perturbations then one can set  $\{c_{l,k} = 1\}$ . Formally, the vector  $\mathbf{x}_\Upsilon$ , looks like that for the EGF, but EAE will have many fewer components, as explained above. In the EAE formulation, each subvector  $\mathbf{x}_l$  corresponds to the coefficients associated with aberration  $A_l^u$ . So,

$$\mathbf{x}_\Upsilon^T = [\mathbf{x}_0, \dots, \mathbf{x}_l, \dots, \mathbf{x}_{N-1}]^T, \text{ and} \quad (37)$$

$$\mathbf{x}_l = [c_{l,0}, \dots, c_{l,k}, \dots, c_{l,M_l-1}, d_{l,0}, \dots, d_{l,k}, \dots, d_{l,M_l-1}]^T, \quad (38)$$

To calculate the elements of the  $\mathbf{H}_\Upsilon$  matrix associated with EAE, Eq. (36) is used in Eq. (35), which, in turn, is substituted in to Eq. (16).

The resulting linear system has the form:

$$\begin{bmatrix} \mathbf{y}_0 \\ \vdots \\ \mathbf{y}_l \\ \vdots \\ \mathbf{y}_{N-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{\Upsilon,0} & \mathbf{H}_{g,0} & \mathbf{H}_{p,0} \\ \vdots & \vdots & \vdots \\ \mathbf{H}_{\Upsilon,l} & \mathbf{H}_{g,l} & \mathbf{H}_{p,l} \\ \vdots & \vdots & \vdots \\ \mathbf{H}_{\Upsilon,N-1} & \mathbf{H}_{g,N-1} & \mathbf{H}_{p,N-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_\Upsilon \\ \mathbf{x}_g \\ \mathbf{x}_p \end{bmatrix}, \quad (39)$$

which, unlike Eq.(29), is a dense linear system that has no obvious structure to exploit.

#### 4.2.1 The Role of Equivalent Aberrations in the EAE

As mentioned above, it may be permissible to reduce the number of planes included in the EAE in Eq. (35), which reduces the size of  $\mathbf{x}_T$ . In short, under the approximations of geometrical optics, if aberrations are present in one or more planes that are conjugate to plane  $l$ , then these conjugate planes may be omitted from the EAE in Eq. (35), effectively replacing the aberrations in these planes with an *equivalent aberration* in plane  $l$ . Demonstrating this requires the concept of equivalent aberrations, which was explained in Sec. 4.1 of [4], but it is included here due to its direct relevance.

Consider the conceptual diagram in Fig. 2, which represents three transmitting surfaces, downstream of some DM given the command vector  $\mathbf{m}$ , that interact with optical radiation that is propagating from top to the bottom in the picture. The light first interacts with surface  $S_0$ , then propagates to the surface  $S_1$ , interacts with it, and finally propagates to  $S_2$ . Let coordinate vectors in the surfaces  $S_0$ ,  $S_1$  and  $S_2$  be given by  $\mathbf{r}_0$ ,  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , respectively. The field just before interacting with  $S_0$  is denoted by  $u_0(\mathbf{r}_0, \mathbf{m})$ , and similarly for  $u_1(\mathbf{r}_1, \mathbf{m})$  and  $u_2(\mathbf{r}_2, \mathbf{m})$ . The relationship between  $u_0(\mathbf{r}_0, \mathbf{m})$  and  $u_1(\mathbf{r}_1, \mathbf{m})$  is defined by the propagation operator  $\Upsilon_{1,0}(\mathbf{r}_1, \mathbf{r}_0)$ , with a similar meaning for the operator  $\Upsilon_{2,1}(\mathbf{r}_2, \mathbf{r}_1)$ , so that

$$u_1(\mathbf{r}_1, \mathbf{m}) = \Upsilon_{1,0}(\mathbf{r}_1, \mathbf{r}_0) u_0(\mathbf{r}_0, \mathbf{m}), \text{ and} \quad (40)$$

$$u_2(\mathbf{r}_2, \mathbf{m}) = \Upsilon_{2,1}(\mathbf{r}_2, \mathbf{r}_1) u_1(\mathbf{r}_1, \mathbf{m}). \quad (41)$$

Consider a thought experiment in which the propagators  $\Upsilon_{1,0}$  and  $\Upsilon_{2,1}$  include the phase aberrations imparted by the surfaces  $S_0$  and  $S_1$ , denoted by  $\phi_0(\mathbf{r}_0)$  and  $\phi_1(\mathbf{r}_1)$ , respectively. Then two propagators can be written as

$$\Upsilon_{1,0}(\mathbf{r}_1, \mathbf{r}_0) = \Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) \exp[j\phi_0(\mathbf{r}_0)] \text{ and} \quad (42)$$

$$\Upsilon_{2,1}(\mathbf{r}_2, \mathbf{r}_1) = \Upsilon_{2,1}^k(\mathbf{r}_2, \mathbf{r}_1) \exp[j\phi_1(\mathbf{r}_1)], \quad (43)$$

where  $\Upsilon_{1,0}^k$  and  $\Upsilon_{2,1}^k$  are known propagators. Then, using Eqs. (42) and (43) inside Eqs. (40) and (41),

$$u_2(\mathbf{r}_2, \mathbf{m}) = \Upsilon_{2,1}^k(\mathbf{r}_2, \mathbf{r}_1) \exp[j\phi_1(\mathbf{r}_1)] \Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) \exp[j\phi_0(\mathbf{r}_0)] u_0(\mathbf{r}_0, \mathbf{m}). \quad (44)$$

Now, suppose one wishes to treat the cumulative effect of the aberrations on  $S_0$  and  $S_1$  with some equivalent aberration only on  $S_1$ , and let this equivalent aberration be represented by  $\phi'_1$ . In other words, does there exist some  $\phi'_1$  that satisfies the condition

$$\begin{aligned} u_2(\mathbf{r}_2, \mathbf{m}) = & \\ & \Upsilon_{2,1}^k(\mathbf{r}_2, \mathbf{r}_1) \exp[j\phi_1(\mathbf{r}_1)] \Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) \exp[j\phi_0(\mathbf{r}_0)] u_0(\mathbf{r}_0, \mathbf{m}) = \\ & \Upsilon_{2,1}^k(\mathbf{r}_2, \mathbf{r}_1) \exp[j\phi'_1(\mathbf{r}_1, \mathbf{m})] \Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) u_0(\mathbf{r}_0, \mathbf{m}) ? \end{aligned} \quad (45)$$

Clearly, the validity of Eq. (45) is independent of the leftmost operator  $\Upsilon_{2,1}^k$ , and after dropping it,  $\phi'_1(\mathbf{r}_1, \mathbf{m})$  must be given by:

$$\exp[j\phi'_1(\mathbf{r}_1, \mathbf{m})] = \frac{\exp[j\phi_1(\mathbf{r}_1)] \Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) \exp[j\phi_0(\mathbf{r}_0)] u_0(\mathbf{r}_0, \mathbf{m})}{\Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) u_0(\mathbf{r}_0, \mathbf{m})}. \quad (46)$$

Ignoring the possibility of zeros in the denominator, from Eq. (46) one can see that there does indeed exist an equivalent aberration  $\phi'_1$  on  $S_1$  that replaces both of the original aberrations on  $S_0$  and  $S_1$ , however, *the price of this substitution is that the equivalent aberration depends on the DM configuration ( $\mathbf{m}$ ), even though the original aberrations do not.* Thus, the value of  $\phi'_1(\mathbf{r}_1, \mathbf{m})$  will fluctuate as the DM modulates  $u_0(\mathbf{r}_0, \mathbf{m})$ . However, there is (at least) one non-trivial choice for the operator  $\Upsilon_{1,0}^k$  that removes the problem of dependence on the DM configuration. Suppose that surfaces  $S_1$  and  $S_0$  are in conjugate planes under geometrical optics approximations. In that case, the propagator  $\Upsilon_{1,0}$  is given by  $\Upsilon_{1,0}^k(\mathbf{r}_1, \mathbf{r}_0) = \delta(\mathbf{r}_0 - \beta\mathbf{r}_1)$ , where  $\delta(\cdot)$  is the Dirac delta, and  $\beta$  is some scalar magnification factor, and Eq. (46) becomes (integrating over  $\mathbf{r}_0$ ):

$$\exp[j\phi'_1(\mathbf{r}_1)] = \exp[j\phi_1(\mathbf{r}_1) + \phi_0(\beta\mathbf{r}_1)], \quad (47)$$

where the second argument of  $\phi'_1(\mathbf{m})$  has been dropped since the  $\mathbf{m}$ -dependent factors cancel, as it depends only the original aberrations  $\phi_0$  and  $\phi_1$ . *Thus, when two planes are conjugate under geometrical optics, their aberrations can be replaced by an equivalent aberration that is independent of the incident field.* On the other hand, Eq. (47) says that, under geometrical optics approximations, aberrations in certain optical systems with many surfaces possibly can be represented by equivalent aberrations in only several planes, if the aberrations in many planes are all conjugate to only a few planes. Solving for the equivalent aberrations under this approximation is a form of tomography and in the future it may be called “aberration tomography.” One should also be careful before assuming two planes are conjugate in a given optical system, especially as certain optics components such as DMs and coronagraphs can easily break conjugacy relationships.

## 5. RELATIONSHIP TO PREVIOUS FPWS FORMULATIONS

The developments presented here allow one to understand the initial assumption in the original paper in 1995 by Malbet et al. [1] that has since been repeated in all subsequent work [e.g., 2, 14, 15]. Under the assumption that  $\Upsilon_{c,d}^u = 0$ , i.e., that there is no unknown aberration downstream of the DM, the 2nd and 3rd terms of Eq. (12) are zero. Then, using Eqs. (8), (9) and (12) one obtains:

$$u_c(\mathbf{r}_c, \mathbf{m}) \text{ “=” } \Upsilon_{c,d}^k(\mathbf{r}_c, \mathbf{r}_d) \exp[j\zeta(\mathbf{r}_d, \mathbf{m})] \{ \Upsilon_{d,0}^k(\mathbf{r}_d, \mathbf{r}_0) u_{*0}(\mathbf{r}_0) + g^u(\mathbf{r}_d) \}, \quad (48)$$

which is the equivalent of, e.g., Eq. 1 of [1], Eq. 2 of [14] or Eq. 1 of [2]. In Eq. (48), the equals sign is shown in quotation marks to emphasize the invalidity of the  $\Upsilon_{c,d}^u = 0$  assumption.

Interestingly, even if an aberration in the post-DM optics is in a plane that is conjugate to the DM-plane or some plane in the pre-DM optics, it cannot be taken into account with the effective pre-DM aberration,  $g^u(\mathbf{r}_d)$  in Eq. (11). Thus, even in this favorable situation, Eq. (48) is not valid. To demonstrate this fact, as a thought experiment, consider the following form for  $\Upsilon_{c,d}^u$ :

$$\Upsilon_{c,d}^u(\mathbf{r}_c, \mathbf{r}_d) = \Upsilon_{c,q}^k(\mathbf{r}_c, \mathbf{r}_q) \exp[j\vartheta_q(\mathbf{r}_q)] \delta(\mathbf{r}_d - \mathbf{r}_q), \quad (49)$$

where  $\vartheta_q(\mathbf{r}_q)$  is the sole unknown aberration downstream of the DM in some plane denoted by index  $q$ , and  $\Upsilon_{c,q}^k$  is a known propagator. The optical conjugacy between the  $d$  and  $q$  planes is expressed by using the Dirac delta function as the propagation operator between these two planes. Then, using Eq. (49), the 2nd term in Eq. (12) becomes (the 3rd term is product of two presumably small unknowns):

$$\Upsilon_{c,q}^k(\mathbf{r}_c, \mathbf{r}_q) \exp[j\vartheta_q(\mathbf{r}_q) + \zeta(\mathbf{r}_q, \mathbf{m})] \Upsilon_{d,0}^k(\mathbf{r}_q, \mathbf{r}_0) u_{*0}(\mathbf{r}_0), \quad (50)$$

which cannot be included in the  $g^u(\mathbf{r}_d)$  function defined in Eq. (11) since the leftmost operator,  $\Upsilon_{c,q}^k$ , integrates over the  $q$ -plane, not the  $d$ -plane as does Eq. (48).

The reader may wonder how various groups have successfully created dark regions when starting with the incorrect Eq. (48). As explained, e.g., in [2], these EFC methods do not attempt to estimate the unknown aberration (essentially  $g^u$  in those formulations), rather they directly estimate the field in the image plane using DM “probes,” under the assumption that the probe deformations are small enough that the effect of the unknown aberration on the changes in the field induced by the DM steps are negligible (which is certainly true when steps are small enough). Thus, as the DM shape converges to the optimal configuration, the field in the image plane is constantly reassessed. One of major difficulties with this paradigm is that as the DM converges to the optimal configuration, the dark hole becomes very dark indeed, making photon counting noise (shot noise) a limiting factor in the accuracy of the technique. *In contrast, the methods presented here make no stipulation that the DM step size is small, which allows the regression to be informed by DM shapes that are relatively far from a dark hole configuration. Then, after the regression procedure has estimated the aberrations, the DM can be commanded to take its optimal configuration.* Another difficulty with current methodology is that after determining the field in the image plane, the DM is commanded to take the configuration that algorithm, be it stroke-minimization, electric field conjugation, or something else, calculates to be optimal for creating the dark hole. However, there is no guarantee that this DM step will be small enough so that the effect of unknown aberration will be negligible. *In contrast, the method presented would allow these algorithms to include the effect of the aberrations whatever the step-size, making solution more accurate. This could improve convergence to the dark hole and possibly make it even darker.*

## 6. CONCLUSIONS

The FPWS methods presented here are potentially exciting because they mitigate two important difficulties with current methodologies, namely the shot-noise limitations in the dark hole (because the regression can be informed by DM configurations that are relatively far from dark solution), and the accuracy of the optimum that is found by the minimization algorithm (because the cost functions will be accurate with larger step-sizes). These methods also have the capability of being used iteratively to help mitigate the effects of the linearization in the regression. For example, once the aberrations have been estimated, they can be included in the known propagators, so that the new unknown propagators, determined by another regression, treat only smaller aberrations.

This article has completely neglected polarization effects, which are likely to be important, since partial linear and circular polarization is caused by reflection off of mirrors. This is critical because orthogonal polarizations do not interfere, so that the diffraction pattern from beam with spatially variable polarization will be different from that derived in the scalar approximation, in which of the light interferes. The effects of polarization in telescopes have been described by [16], who shows that a telescope is better described by a point-spread-matrix (which operates on the Stokes vector) than a scalar PSF. The regression methods described here can be made fully polarimetric using methods similar to those in [4] and in [13], which presented the vector generalization of the empirical Green's function called the empirical Green's tensor. The author plans to present the polarimetric versions of these equations in a followup article. Another limitation of the proposed methods is the accuracy and reproducibility of the DM shape  $\zeta(\mathbf{m})$ , which leads into DM calibration issues that are beyond the scope of this discussion. Whatever the achievable accuracy of the DM, in the end, the true value of  $\zeta(\mathbf{m})$  is a stochastic quantity, and it would be necessary to characterize some of its statistical properties in order to assess the likely consequences of DM figure error. Demonstrating the suitability of the methods presented here for a space mission, first in simulation and then in the laboratory, is left to future efforts.

## Acknowledgments

The author thanks Jim Breckinridge for providing comments and engaging in helpful discussions. This work has been supported by NSF Award #1600138 to the University of Michigan.

## REFERENCES

- [1] Malbet, F., Yu, J. W., and Shao, M., “High-Dynamic-Range Imaging Using a Deformable Mirror for Space Coronagraphy,” *PASP* **107**, 386 (Apr. 1995).
- [2] Groff, T. D., Eldorado Riggs, A. J., Kern, B., and Jeremy Kasdin, N., “Methods and limitations of focal plane sensing, estimation, and control in high-contrast imaging,” *Journal of Astronomical Telescopes, Instruments, and Systems* **2**, 011009 (Jan. 2016).
- [3] Tang, H., Rud, M., Demers, R., Goullioud, R., Krist, J., Shi, F., and Zhao, F., “The WFIRST/AFTA coronagraph instrument optical design,” in [*Techniques and Instrumentation for Detection of Exoplanets VII*], Proc. SPIE **9605**, 960504 (Sept. 2015).
- [4] Frazin, R. A., “Statistical framework for the utilization of simultaneous pupil plane and focal plane telemetry for exoplanet imaging. I. Accounting for aberrations in multiple planes,” *Journal of the Optical Society of America A* **33**, 712 (Apr. 2016).
- [5] Born, M. and Wolf, E., [*Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, seventh expanded edition*], The Press Syndicate of the University of Cambridge (1999).
- [6] Goodman, J. W., [*Statistical Optics, 2nd Edition*], John Wiley and Sons, Inc. (2015).
- [7] Groff, T. D. and Kasdin, N. J., “Kalman filtering techniques for focal plane electric field estimation,” *J. Opt. Soc. Am. A* **30**, 128–139 (Jan 2013).
- [8] Frazin, R. A., “Utilization of the Wavefront Sensor and Short-exposure Images for Simultaneous Estimation of Quasi-static Aberration and Exoplanet Intensity,” *ApJ* **767**, 21 (Apr. 2013).
- [9] Fathipour, V., Memis, O. G., Jang, S. J., Khalid, F., Brown, R. L., Hassaninia, I., Gelfand, R., and Mohseni, H., “Isolated nanoinjection photo detectors for high-speed and high-sensitivity single-photon detection,” in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **8868**, 3 (Sept. 2013).

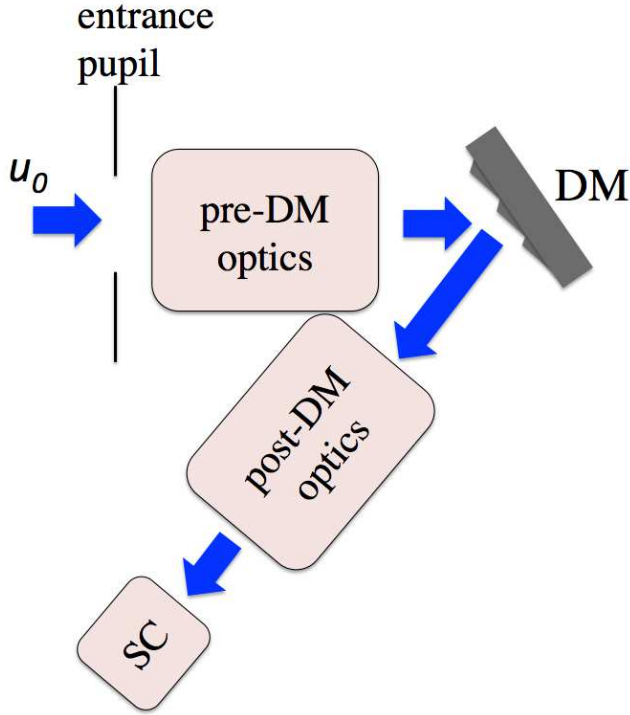


Figure 1. Schematic diagram showing the initial field  $u_0$  impinging on the telescope entrance pupil, passing through the pre-DM optical train, reflecting off the deformable mirror (DM), passing through the post-DM optical train and, finally, forming an image in the science camera (SC). The arrows indicate the direction of light propagation.

- [10] Finger, G., Baker, I., Alvarez, D., Ives, D., Mehrgan, L., Meyer, M., Stegmeier, J., Thorne, P., and Weller, H. J., “Evaluation and optimization of NIR HgCdTe avalanche photodiode arrays for adaptive optics and interferometry,” in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **8453** (July 2012).
- [11] Finger, G., Baker, I., Alvarez, D., Ives, D., Mehrgan, L., Meyer, M., Stegmeier, J., and Weller, H. J., “SAPHIRA detector for infrared wavefront sensing,” in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **9148**, 17 (Aug. 2014).
- [12] Mazin, B., “MKIDs for TMT,” in [*TMT in the Astronomical Landscape of the 2020s, Thirty Meter Telescope Science Forum, held 16-19 July, 2014 in Tucson Arizona. Online at: <http://conference.ipac.caltech.edu/tmtsf2014/>*], (July 2014).
- [13] Frazin, R. A., “Empirical Green’s Function Approach for Utilizing Millisecond Focal and Pupil Plane Telemetry in Exoplanet Imaging,” in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **9909** (2016).
- [14] Bordé, P. J. and Traub, W. A., “High-Contrast Imaging from Space: Speckle Nulling in a Low-Aberration Regime,” *ApJ* **638**, 488–498 (Feb. 2006).
- [15] Martinache, F., Guyon, O., Clergeon, C., and Blain, C., “Speckle Control with a Remapped-Pupil PIAA Coronagraph,” *PASP* **124**, 1288–1294 (Dec. 2012).
- [16] Breckinridge, J. B., Lam, W. S. T., and Chipman, R. A., “Polarization Aberrations in Astronomical Telescopes: The Point Spread Function,” *PASP* **127**, 445–468 (May 2015).

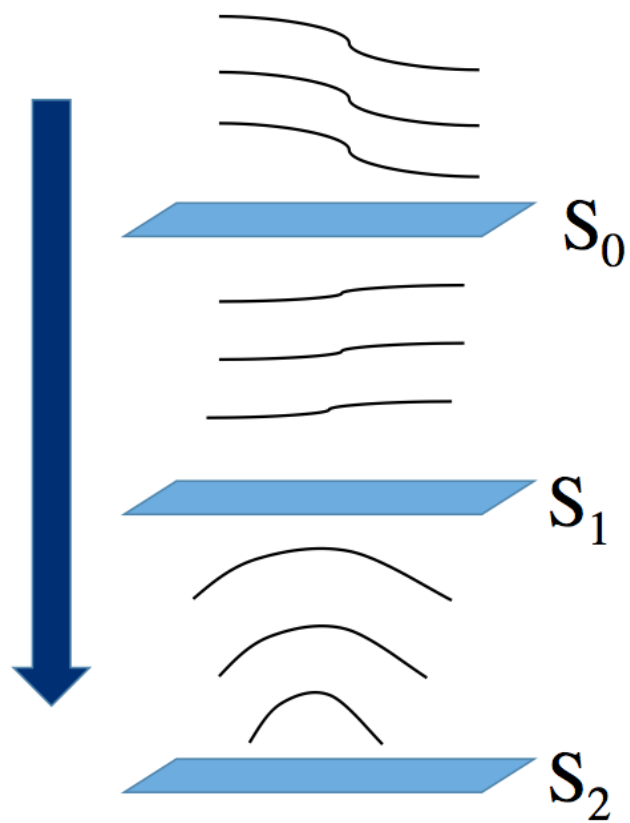


Figure 2. Conceptual diagram depicting light passing through phase screens on surfaces  $S_0$  and  $S_1$  before arriving at surface  $S_2$ . The arrow indicates the direction of propagation, and the black curves indicate a sequence of wavefronts.